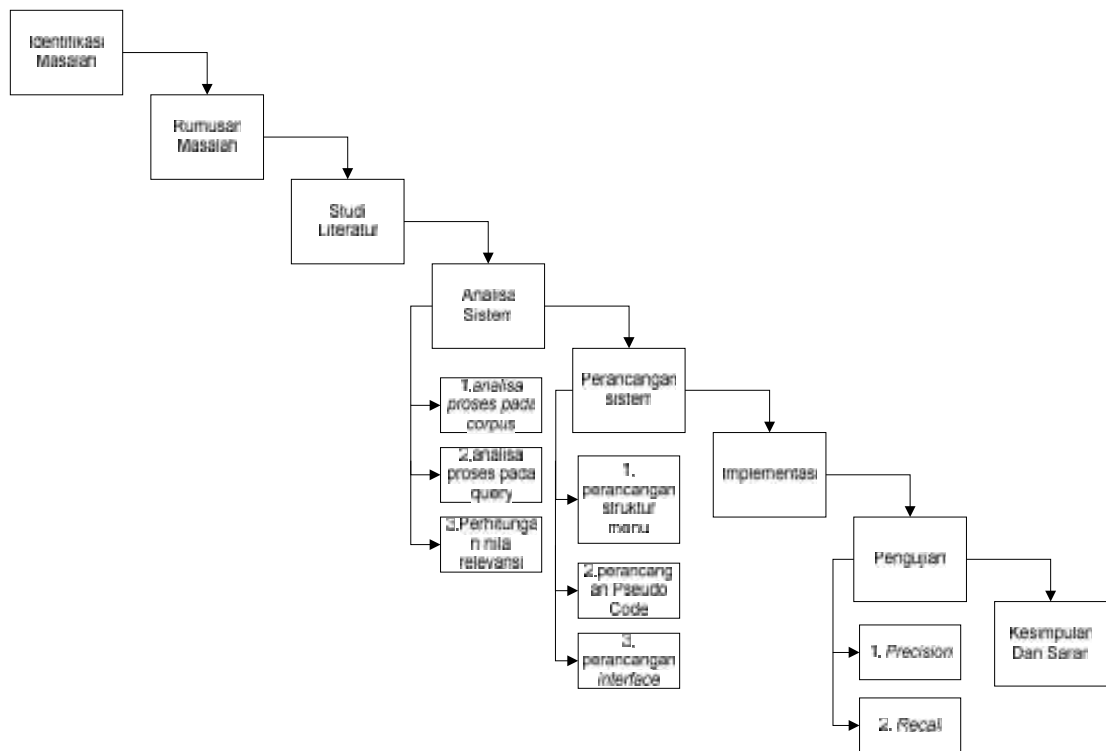


BAB III

METODOLOGI PENELITIAN

Metodologi penelitian merupakan sistematika tahap-tahap yang dilaksanakan dalam pembuatan tugas akhir. Adapun tahapan yang dilalui dalam pelaksanaan penelitian ini adalah sebagai berikut:



Gambar 3.1 flowchart Penyusunan Tugas Akhir

3.1 Identifikasi Masalah

Pada tahapan ini dilakukan identifikasi permasalahan mengenai pentingnya bagi *user* untuk memperoleh dokumen ayat Al-Qur'an dan terjemahan yang relevan sesuai dengan kebutuhannya dan masukan *query* dari sekumpulan informasi (dokumen).

3.2 Perumusan Masalah

Dari tahapan identifikasi masalah, didapatkan rumusan masalah tentang bagaimana proses membangun aplikasi pencarian ayat Al-Qur'an berdasarkan terjemahan bahasa Inggris dengan konsep *Information Retrieval* dengan Porter *Stemmer Enhancement* menggunakan model ruang vektor .

3.3 Studi Literatur

Studi literatur dilakukan dengan mempelajari konsep-konsep yang berkaitan dengan penelitian ini, seperti sistem temu balik informasi, model ruang vektor, tahapan *text operations/preprocessing* seperti penghapusan *markup*, penghapusan *stopwords*, *stemming*, *query operations*, dan *weighting* melalui literatur-literatur seperti buku, jurnal, dan sumber ilmiah lain seperti situs internet ataupun artikel dokumen teks yang berhubungan.

3.4 Analisa Sistem

Pada tahapan ini, ditentukan analisa masukan (*input*) untuk membangun sistem, analisa proses pada sistem, serta tahapan pembangunan *information retrieval system* sebelum diterapkan Model Ruang Vektor dalam pengukuran kesamaan (*similarity measure*). Secara garis besar tahapan pembangunan sistem pencarian ayat Al-Qur'an dalam tugas akhir ini adalah:

1. *Text Operation*

Pertama, mengumpulkan dokumen yang akan diindeks (korpus). Koleksi dokumen (korpus) yang digunakan dalam membangun sistem adalah korpus terjemahan Al-Qur'an dalam bahasa Inggris dan korpus gambar (tulisan arab) ayat Al-Qur'an berformat *.png. Dari dua koleksi dokumen tadi dilakukan proses untuk penyusunan kembali korpus untuk membentuk sebuah korpus baru berisi kumpulan subtopik Al-Qur'an. Satu dokumen subtopik mewakili dari beberapa ayat Al-Qur'an. Pembagian subtopik merujuk dari subtopik Al-Qur'an terjemahan bahasa Indonesia oleh Wulandari, 2012. Hal ini dilakukan untuk menghindari salah tafsir dari pengguna. Jika dijadikan satu ayat sebagai satu dokumen, ditinjau dari terjemahannya, banyak ayat pendek justru tidak selesai kalimatnya, atau tidak mengandung suatu makna yang jelas. Korpus

subtopik Al-Qur'an nantinya akan digunakan sebagai bahan indeks dokumen untuk melakukan tahapan *indexing*.

a. Penghapusan *Markup*

Pemrosesan pemisahan rangkaian kata pada teks terjemahan Al-Qur'an pada tahapan *preprocessing* adalah proses memisahkan kalimat menjadi potongan kata/*term*. Pada proses ini juga menghilangkan karakter tanda baca (*stoplist*) dan mengubah kata ke dalam huruf kecil.

b. Penghapusan *Stopwords*

Setelah proses pemisahan rangkaian kata, dilakukan proses penghapusan *stopwords* untuk kata-kata yang sering muncul dan dianggap tidak mencerminkan isi dari dokumen. Dalam bahasa Inggris yang merupakan bagian dari *stopwords* adalah artikel(*articles*), kata depan(*prepositions*), dan kata penghubung(*conjunctions*). Dengan membuang *stopwords*, ukuran struktur *indexing* dapat dikurangi dan dapat mengurangi *recall* dokumen yang tidak sesuai.

c. *Stemming*

Proses *stemming* dilakukan untuk menjadikan kata yang ke bentuk dasar, dengan menghilangkan awalan atau akhiran dari sebuah kata. *Stemming* di sini menggunakan kamus kata dasar sebagai pembanding kata-kata yang ada dalam terjemahan Al-Qur'an yang telah melalui tahapan *stemming*. Selain menggunakan Porter *Stemmer* sebagai algoritma untuk memotong akhiran kata, juga akan dilanjutkan dengan pengecekan dan pemotongan awalan kata dengan mengadaptasi aturan algoritma Nazief dan Adriani untuk peningkatan kualitas *stemming* untuk memperoleh hasil yang lebih baik.

2. *Indexing*

Pada saat *logical view* di tahap *text operations* telah ditetapkan, *database manager* (menggunakan *Database manager module*) akan membangun indeks dari teks. Sebuah indeks adalah sebuah struktur data kritis karena mengizinkan pencarian cepat terhadap data berukuran besar. Struktur-struktur indeks yang berbeda dapat digunakan, tetapi yang paling sering adalah *inverted file*. Sumber-sumber (waktu dan tempat penyimpanan) yang terpakai

untuk mendefinisikan teks *database* dan membangun indeks diselesaikan dengan melakukan pemrosesan *query* (*querying*) terhadap IR berulang kali.

3. *Query Operation*

Query operations mencakup pemberian bobot dan perluasan *query*. *User* pertama kali harus menentukan sebuah kebutuhan informasi yang kemudian akan diuraikan (*parsed*) dan diubah oleh operasi teks yang sama yang diterapkan pada teks dokumen. Lalu, operasi-operasi *query* bisa diterapkan sebelum *query* yang sebenarnya, yang menyediakan sebuah representasi sistem sehingga dapat menghasilkan kebutuhan *user*. *Query* kemudian diproses untuk memperoleh dokumen-dokumen yang dikembalikan atau dicari (*retrieved documents*). Pemrosesan *query* yang cepat dimungkinkan dengan struktur indeks yang sebelumnya sudah dibangun.

4. *Searching & ranking*

Setelah menghitung pembobotan, dilakukan perhitungan kemiripan (*similarity*) antara dokumen dengan *query* dengan menerapkan pemodelan ruang vektor. Melakukan analisis terhadap persamaan model ruang vektor untuk mengetahui nilai korelevanan/kemiripan (*similarity*) dokumen terhadap *query* sehingga dokumen dapat diurutkan (*diranking*). Setelah dokumen *diranking*, sejumlah tetap dokumen *top-scoring* dikembalikan kepada pengguna.

3.5 Perancangan Sistem

Pada tahapan ini akan dilakukan analisa pembangunan *information retrieval system* terhadap kebutuhan pengguna dan kebutuhan perangkat lunak, dijelaskan secara rinci tentang proses dari sistem temu kembali informasi yang akan dibangun sehingga mempermudah pemahaman terhadap sistem.

Tahap perancangan *information retrieval system* dengan model ruang vektor merupakan tahapan dalam membuat rincian sistem temu balik informasi berdasarkan analisa agar dapat dimengerti oleh pengguna. Tiga rancangan utama yang akan dilakukan, yaitu:

1. Perancangan sistem temu kembali ayat Al-Qur'an yang akan dibangun menggunakan pemodelan berorientasi objek, yaitu OOAD (*Object Oriented Analysis and Design*).
2. Transformasi koleksi dokumen ke dalam bentuk *inverted index database*.
3. Perancangan antarmuka sistem (*interface*) yang baik sehingga mudah digunakan (*user friendly*).

3.6 Implementasi Sistem

Pada implementasi sistem akan dilakukan pembuatan modul-modul yang telah dirancang dalam tahap perancangan ke dalam bahasa pemrograman. Berikut adalah spesifikasi lingkungan implementasi perangkat keras dan perangkat lunak yang akan digunakan:

1. Perangkat keras

<i>Processor</i>	:	<i>Intel(R) Core(TM) i3 CPU M 350, 2.3GHz</i>
<i>Memory (RAM)</i>	:	2 GB
<i>Harddisk</i>	:	320 GB

2. Perangkat lunak

Sistem Operasi	:	<i>Windows 7 Professional</i>
Bahasa Pemrograman	:	PHP
<i>Tools</i> Perancangan	:	<i>Notepad++</i>

3.7 Pengujian Sistem

Pengujian merupakan tahapan sistem yang akan dijalankan. Tahap pengujian diperlukan untuk menjadi ukuran bahwa sistem dapat dijalankan sesuai dengan tujuan. Pengujian sistem temu balik informasi dalam tugas akhir ini dilakukan dengan cara mengukur kualitas *text retrieval*. Ukuran yang digunakan untuk mengukur kualitas dari *documents retrieval* adalah *precision* dan *recall* berdasarkan hasil perhitungan kemiripan (*similarity measure*).

3.8 Kesimpulan dan Saran

Tahapan ini akan membahas tentang kesimpulan yang dihasilkan dari penelitian tugas akhir, kesimpulan diambil dari proses analisa kepada implementasi dan pengujian. Pada tahapan saran, penelitian ini diharapkan dapat memberikan sebuah catatan rekomendasi untuk menyempurnakan dan mengembangkan penelitian sistem temu kembali pencarian ayat Al-Qur'an.